

## 主旋律の類似性について

門田 隆史 石野 明 竹田 正幸 松尾 文碩

九州大学大学院システム情報科学研究科

**要旨.** 近年、デジタル情報として蓄積される音楽情報が急増しており、その膨大な音楽情報に対する検索システムの開発や、計算機による音楽分析などが注目されつつある。そこで最も重要な問題のひとつとして、楽曲の類似性をいかに定義するかという問題があげられる。Mongeauら(1990)は、楽曲間の非類似度を、楽譜上の音符と休符の列に対する重み付き編集距離としてあたえている。この指標は、同一の変奏曲内において主題と各変奏の近さを順序づける際には有効であるが、複数の変奏曲に対して適用すると問題が生じる。そこで本論文では、新たな類似性指標を提案する。楽曲の類似には、様々な要因が考えられるが、ここでは、(1)音程、(2)リズム、の二つに着目し、それぞれの観点から指標をあたえ、その検証を行なった。その結果、(1)の観点からあたえた指標については、上述の複数の変奏曲に対する順序づけにおいて、Mongeauらの指標より有効であることを示した。また、(2)の観点からあたえた指標については、特に歌謡曲データにおいて、(1)の観点では得られなかった類似フレーズを抽出できることが判明した。

## On Melodic Similarity

Takashi Kadota Akira Ishino Masayuki Takeda Fumihiko Matsuo

Department of Informatics, Kyushu University 33, Fukuoka 812-8581, JAPAN

**Abstract.** Any monophonic score can be regarded as a series of ordered pairs with the pitch of the note and its length. Mongeau and Sankoff (1990) adapted concepts from the theory of sequence comparison to define the overall dissimilarity between two monophonic scores. Their measure is effective in the sense that it arranges a set of variations on a theme by Mozart in a reasonable order which coincides with subjective impressions. However, our experimental results reveal that a problem arises when dealing with the variations on several themes. In this paper we propose new similarity measures to cope with this problem. Similarity on musical scores can be defined from various viewpoints, and here we focus on the two points: (1) pitch and (2) rhythm. We provide similarity measures from both points, and try to verify their effectiveness. The results show that the pitch based measures are better than the one due to Mongeau and Sankoff, and that the rhythm based measures find similar phrases in popular songs which cannot be found by the pitch based measures.

### 1 はじめに

通信カラオケの普及やインターネットを介した音楽配信ビジネスなどにみられるように、ディ

ジタル情報として蓄積される音楽情報が急増しており、音楽情報検索システムの開発や、計算機による音楽分析などが注目されつつある。このように計算機で音楽情報を扱う際、重要な問

題の一つに、楽曲の類似性をいかに定義するか、という問題がある。本論文では、主旋律の類似性の問題を扱う。

音楽情報データベースの検索においては、メロディの断片を質問とした検索を行なえることが望ましい。作曲家や音楽家など、音楽の専門家をユーザとした場合には、質問は、キーボード等を用いて正確な形で与えられるものと考えて良い。一方、一般ユーザには、質問を正確な形で用意することが難しいため、いくつかのシステムでは、ハミングや歌声による検索を行なえるようにしている [2, 3, 1]。一般に、音楽演奏の音響信号を楽譜レベルの情報に変換することは難しいが、歌声やハミングによる入力では、さらに、入力された旋律情報の精度が低いという問題が生じる。この問題に対しては、メロディを、コンツア(contour)、すなわち、音程が上がったか、下がったか、同じ高さか、の3段階の粗い基準によって表現した記号列に変換する方法が用いられる。

本論文では、音楽演奏の音響信号を直接扱うのではなく、楽譜レベルの情報を対象として主旋律の類似性の問題を考える。Mongeauら [4] は、音楽分析や著作権問題への応用を目的として、楽譜レベルでの情報を対象とした類似性指標を提案している。Mongeauらは、楽譜上の音符(または休符)の列を対象とした文字列間の重み付き編集距離(weighted edit distance)として、楽曲間の非類似度を定義し、これを、モーツァルトの『きらきら星変奏曲』における主題と各変奏の近さの順序づけの問題に適用し、直感に合った結果が得られることを報告している。しかし、別の変奏曲を混ぜた上で、改めて順序づけを行なうと、本来類似していない、別の変奏曲の主題や変奏が、高い順位を得ることが判明した。したがって、類似曲検索に用いることはできない。そこで本論文では、新たな類似性指標を提案する。

楽曲の類似性を扱う場合、一定のまとまりをもった音列(フレーズ)を単位として対比させた上で、それらの対比を総合して判断すべきである。そこで、まず、フレーズに対する類似性指標を定義し、次に、その指標をもとに楽曲間の

類似性指標を定めることにした。

楽曲の類似には、様々な要因が考えられるが、本論文では、そのうち、(1)音程、(2)リズム、の二つの観点に絞って類似性を考える。(1)の観点からあたえた指標については、上に述べた複数の変奏曲データを用いて、Mongeauらの指標との比較検証を行なった結果、その有効性を確認できた。また、(2)の観点から得た指標については、(1)では得られなかった類似フレーズを抽出できることが確認できた。

## 2 Mongeauらの非類似性指標

本章では、まず、Mongeauら [4] による類似性指標の概略を述べ、その問題点を考察する。

### 2.1 指標の概略

Mongeauら [4] は、楽曲の旋律を、音符や休符を1文字とする文字列とみなし、文字列間の重み付き編集距離(weighted edit distance)として楽曲間の非類似度を定義した。編集操作(edit operation)としては、基本的に、音符の置換、挿入、削除の3種類を考える。一つの音符は、音程と音長の対であるから、各置換操作のコスト(重み)は、置換に関与する二つの音符の音程と音長に依存して定める。音程に関しては和声学に基づく協和・不協和の関係を用いてコストを定め、音長については音長差の絶対値をコストとし、二つの量の加重和を置換操作のコストとする。なお、休符は、音程のない特別な音符として扱う。また、挿入・削除操作に関わるコストは、音長が0の同じ音程の音符との置換とみなして定める。さらに、音符の列への独自の編集操作として、1対多、多対1の音符の対応をとる、分割(fragmentation)、統合(consolidation)操作を導入している。分割・統合においては、各音符の音程情報に基づくコストの和と音長の和の差との和をコストとする。

表 1: Mongeau らの指標による結果の一部

順位	K.265-th		K.25-th	
	曲名	非類似度	曲名	非類似度
1	K.265-5	28.4	K.25-7	11.5
2	K.354-4	38.5	K.25-1	25.8
3	K.265-12	39.9	K.25-5	26.8
4	K.25-5	42.5	K.354-4	35.9
5	K.25-th	46.8	K.265-12	36.8
6	K.265-9	47.0	K.265-5	42.4
7	K.354-12	47.2	K.265-th	46.8
8	K.265-11	48.9	K.25-3	49.1
9	K.25-7	49.0	K.354-7	49.6
10	K.354-2	54.6	K.265-11	51.2

## 2.2 検証と考察

Mongeau ら [4] は、モーツァルトの変奏曲 K.265 (『きらきら星変奏曲』) を対象に、各変奏が主題にどれだけ類似しているかの順位づけを行ない、この順位づけが直感的なそれとほぼ一致するという結果を報告している。しかし、K.265 以外の変奏曲に対する結果についてはふれていない。

そこで、まず、別の変奏曲 9 曲に対して同様の検証を行なった。その結果、同一の変奏曲内での順位づけは、いずれの変奏曲の場合も、ほぼ妥当であることが確認できた。

しかし、この指標を類似曲検索に用いるためには、複数の変奏曲を混ぜた中から、同一の主題とその変奏の対を非類似度の値の小さいものとして拾い出すことができなければならない。

そこで、K.265 に上で述べた変奏曲 9 曲を加え、これに対して同様の順位づけによる検証を行なった。表 1 にその結果を示す。なお、簡単のため、ここでは、変奏曲 3 曲 (K.25, K.265, K.345) から主題と変奏を適当に選択したものについての結果のみを示した。ここで、たとえば、「K.265-th」「K.265-5」は、K.265 の主題と第 5 変奏をそれぞれ表すものとする。

この結果において、K.265 と K.25 のそれぞれの主題に対して、同じ曲の変奏よりも、別の変奏曲の変奏が上位に順序づけられている場合がある。これは、ここで示した以外の曲に対する結果でも同様であった。しかし、曲を実際に聴いて判断したところ、やはり、同じ曲の変奏同士が、より似て感じられることが多い。した

がって、Mongeau らの指標を類似曲検索に用いる場合には、この点が問題となることがわかる。

それでは、なぜこれらの類似度が高く評価されたのであろうか。第 1 に、音符の挿入・削除によってタイミングのずれた対応をとっているという問題がある。音楽においてタイミング情報は重要であるから、これらの操作には大きなコストを課すべきである。

第 2 に、音長情報に基づくコストが音長の差のみに依存しており、音長にかかわらず同じコストであるという問題がある。たとえば、音程情報に基づくコストが等しいとき、4 分音符同士の不一致と 16 分音符同士の不一致は同じコストであるが、実際は、より短い音の不一致に低いコストを与えるべきであろう。

第 3 に、音の協和・不協和の関係は感覚によるところが大きく、コストの設定が難しいという問題がある。Mongeau らの指標では、和声学に基づいてコストをあたえている。このコストのあたえ方は、かなり細かいが、具体的な数値に特に根拠があるわけではない。

## 3 主旋律の記号列表現

旋律の記号列表現としては、以下の表現法がよく用いられる。

- (1) 音符または休符の列。各音符は、音程 (pitch) と音長 (duration) の対である。
- (2) 連続した 2 音符間の音程差 (interval) と音長の対を一つの記号とした記号列。
- (3) 上の音程差を上昇、下降、同一、の 3 種類で表したもの。音長を無視した場合、コンツァ (contor) と呼ばれる。

明らかに、(2)(3) は (1) から得ることができる。これらの表現は、いずれも音符を単位としていることに注意されたい。

これに対し、本論文では、各音符を 16 分音符に分割し、音程を表す記号の列として表すことにする。これにより、分割・統合という編集操作は不要となる。また、音程情報については、オクターブの違いは無視し、同じ音程として扱

う。こうすることで、半音ずつ12回平行移動してすべての組合せを考えれば、調性の問題を人手によらずに解決できる。

$\mathcal{P}$  を、音程を表す12個の記号

$$C, C^\sharp (= D^b), D, D^\sharp (= E^b), E, F, F^\sharp (= G^b), G, G^\sharp (= A^b), A, A^\sharp (= B^b), B$$

と、休符を表す特別な記号  $R$  から成る有限集合とする。このとき、(1)の音符・休符列は、 $\mathcal{P}$  を字母とする記号列に変換できる。

しかし、得られた記号列から(1)の記号列は得られない。そこで、各16分音符に、音程とともに、それがもとの音符の先頭であったか否かの情報をもたせる。すなわち、旋律を  $\Sigma = \mathcal{P} \times \{0, 1\}$  を字母とする列

$$\langle a_1, b_1 \rangle, \langle a_2, b_2 \rangle, \dots, \langle a_n, b_n \rangle$$

として表現する。ここに、 $a_i \in \mathcal{P}$ ,  $b_i \in \{0, 1\}$  である。これにより、オクターブの違いを無視すれば、この記号列から(1)の記号列を復元することができる。

本論文では、上記のように記号列表現された旋律を対象とし、その間の類似性について論じる。指標の設計に際しては、この記号列表現を基本とし、必要に応じて、アルファベットインデキシング (alphabet indexing) とよぶ準同型写像により別の字母をもつ記号列に変換する。たとえば、第5章と第6章では、音程を重視した類似性指標について論じるが、この際、記号列  $\langle a_1, b_1 \rangle, \langle a_2, b_2 \rangle, \dots, \langle a_n, b_n \rangle$  を、単に

$$a_1, a_2, \dots, a_n$$

として扱う。一方、第7章においては、リズムを重視した類似性指標について論じる。この際には、音程を表す12個の記号をすべて同一視して記号  $N$  で置き換えることにより、 $\mathcal{P} \times \{0, 1\}$  上の記号列を、

$$\{N, R\} \times \{0, 1\}$$

上の記号列に変換する。この場合、新たに得られた4個の記号は、それぞれ、音符の先頭、音符の継続、休符の先頭、休符の継続を表すことになる。

## 4 類似性指標の統一的枠組

この章では、Tamari ら [5] による文字列間の類似性指標の統一的枠組の概略を述べる。この枠組では、指標を、パターン集合とパターンのスコアを与える関数の対によって表し、二つの文字列間の類似度を、その共通パターンの最大スコアとして定義する。この枠組みは、古典和歌の集積から類似歌を自動抽出するための指標定義の目的で導入されたもので、次のような利点をもつ。

- 代表的な非類似性指標である編集距離およびその変種をすべて表現できること。
- 文字列間の類似性が共通パターンとして陽に与えられるため、指標を直感的に把握しやすいこと。

$\Sigma$  をアルファベットとする。パターンとは  $\Sigma$  上の言語の‘表現’であり、各パターン  $\pi$  に対して、 $\pi$  の表す言語  $L(\pi)$  は一意に定まるものとする。パターン  $\pi$  が文字列  $w \in \Sigma^*$  に合致するとは、 $w \in L(\pi)$  であるときをいう。パターン  $\pi$  が二つの文字列  $x, y \in \Sigma^*$  の共通パターンであるとは、 $\pi$  が両方に合致すること、すなわち、 $x, y \in L(\pi)$  であるときをいう。実数全体の集合を  $\mathbf{R}$  で表す。

$\Sigma$  上の文字列の類似性指標とは、対  $(\Pi, \Phi)$  をいう。ここで、 $\Pi$  はパターン集合といい、各要素  $\pi \in \Pi$  について  $\Sigma$  上の言語  $L(\pi)$  が対応する。また、 $\Phi$  は  $\Pi$  から  $\mathbf{R}$  への関数で、パターンスコア関数とよばれる。類似性指標  $(\Pi, \Phi)$  のもとの文字列  $x, y \in \Sigma^*$  の類似度  $\text{SIM}_{(\Pi, \Phi)}(x, y)$  を、次式で定義する。

$$\text{SIM}_{(\Pi, \Phi)}(x, y) = \max\{\Phi(\pi) \mid x, y \in L(\pi)\}$$

非類似度の場合には、上式において、最大値でなく最小値をとる。

上記の枠組みのもとで類似性指標を設計する際には、以下の二つを決定しなければならない。

- どのような形式のパターンによって類似構造 (共通構造) を表すべきか。
- パターンとして表現された各々の類似構造に、どのようにスコアを割り当てるべきか。





図 1: 指標 A で問題となるフレーズ

ここで、関数  $f: \Pi \rightarrow \mathbf{R}$  は、

$$f(\pi) = \begin{cases} \Phi_1(\pi), & |\pi| \geq s \text{ のとき} \\ 0, & |\pi| < s \text{ のとき} \end{cases}$$

とする。パラメータ  $s$  は連続一致の長さに関する閾値である。

たとえば、パラメータの値を  $s = 2$  とするとき、パターン  $\pi = abb\phi ccdd\phi ee$  においては、分解  $\langle aab, \phi\phi, ccdd, \phi, ee \rangle$  が最大値  $3 + 0 + 4 + 0 + 0 = 7$  をあたえ、これが  $\Phi_2(\pi)$  の値となる。

### 【類似性指標 B】

パターン集合:  $\Pi = (\Sigma \cup \{\phi\})^*$

パターンスコア関数: 定義 2 により定まる写像  $\Phi_2: \Pi \rightarrow \mathbf{R}$ .

## 5.3 類似性指標 C

後述するように、変奏曲による検証により、指標 B は、5.2 節で述べた指標 A の問題点を解消していることが確認できた。しかし、その検証の中で、新たに次のような問題が生じていることがわかった。図 2 に示す例をみてみよう。図に示したフレーズ B1 と B2 は、

ドφドソφソラ...

という共通パターンをもつが、指標 B で、たとえば  $s = 4$  としたとき、類似度は 0 となる。しかし、これらを実際に聴いてみると、十分似たものであった。

そこで、共通パターン中におけるワイルドカード  $\phi$  の連続の長さが閾値  $t$  を超えないときには、音程の一致が継続しているものとみなすことを考えた。

**定義 3** パターンスコア関数  $\Phi_3: \Pi \rightarrow \mathbf{R}$  を以下で定める。

$$\Phi_3(\pi) = \max \left\{ \sum_{i=1}^k g(\pi_i) \mid \begin{array}{l} \langle \pi_1, \dots, \pi_k \rangle \\ \text{は } \pi \text{ の分解} \end{array} \right\}$$

ここで、関数  $g: \Pi \rightarrow \mathbf{R}$  は、

$$g(\pi) = \begin{cases} 0, & \pi \text{ が } \phi^{t+1} \text{ を含むとき} \\ \Phi_2(\pi), & \text{そうでないとき} \end{cases}$$

とする。

### 【類似性指標 C】

パターン集合:  $\Pi = (\Sigma \cup \{\phi\})^*$

パターンスコア関数: 定義 3 により定まる写像  $\Phi_3: \Pi \rightarrow \mathbf{R}$ .

## 5.4 評価

変奏曲データ K.25, K.265, K.354 を用いて、類似性指標 A, B, C の評価を行なった。曲データからひとまとまりのフレーズを切り出すための有効な方法がないので、小節を単位とする  $N$ -gram を用いることとし、人手による検証の行ないやすさから  $N=2$  とした。得られた総フレーズ数は 563 であった。結果の一部を表 2 に示す。ただし、指標 B, C において、 $s = 12, t = 4$  とし

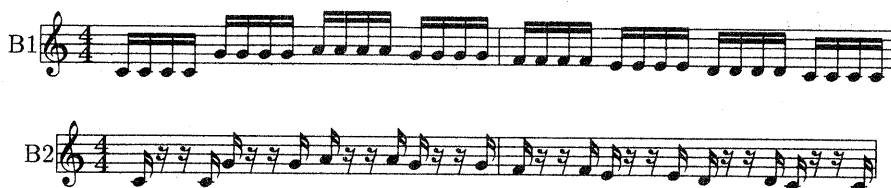


図 2: 指標 B で問題となるフレーズ

表 2: K.265-th(1) に対する各指標の結果

順位	指標 A		指標 B		指標 C	
	フレーズ	類似度	フレーズ	類似度	フレーズ	類似度
1	K.265-th(1)	32.0	K.265-th(1)	32.0	K.265-th(1)	32.0
2	K.265-th(21)	31.0	K.265-th(21)	28.0	K.265-th(21)	31.0
3	K.265-th(5)	31.0	K.265-th(5)	27.0	K.265-th(5)	31.0
4	K.265-th(13)	30.0	K.265-th(13)	20.0	K.265-th(13)	30.0
5	K.265-5(5)	23.0	K.265-th(8)	16.0	K.265-5(5)	23.0
6	K.265-th(3)	22.0	K.265-9(16)	16.0	K.265-5(13)	21.0
7	K.265-th(9)	22.0	K.25-th(16)	13.0	K.265-th(17)	21.0
8	K.265-5(13)	21.0	K.265-9(8)	13.0	K.265-9(21)	19.0
9	K.265-th(7)	21.0	K.25-th(7)	0.0	K.265-9(1)	19.0
10	K.265-th(17)	21.0	K.25-th(8)	0.0	K.265-5(1)	19.0
11	K.265-9(21)	20.0	K.25-th(9)	0.0	K.265-5(21)	19.0
12	K.265-9(5)	19.0	K.25-th(10)	0.0	K.265-9(5)	19.0
13	K.265-9(1)	19.0	K.25-th(11)	0.0	K.265-9(16)	16.0
14	K.265-5(21)	19.0	K.25-th(12)	0.0	K.265-th(8)	16.0
15	K.265-5(1)	19.0	K.25-th(15)	0.0	K.265-11(5)	15.0

た。ここで、たとえば「K.265-th(1)」「K.265-th(21)」は、それぞれ、K.265 主題の第 1 小節から始まる 2 小節のフレーズと第 21 小節から始まる 2 小節のフレーズを表すものとする。

まず、指標 A と指標 B の結果を比較してみよう。これらの結果で注目すべき点は、指標 B によって、多くのフレーズの類似度が減少している点であろう。実際に聴いて判断したところ、類似度の値が大きいものについては、同じ K.265-th(1) の繰り返し部分に当たるフレーズや、極めて似た同曲の変奏であった。一方で、類似度が小さくなったフレーズの多くは実際に似ていなかった。このことから、指標 A から B への変更において目的としていた効果は得られたといえる。

次に、指標 C の結果を見てみよう。ここでは、いくつかのフレーズの類似度の値が再び大きくなっている。これらのフレーズを聴いて判断し

たところ、実際に似たフレーズであった。このように、指標 B によって類似度の値が小さくなったフレーズのうち、比較的「似ている」可能性の高いものに限り、再度、類似度の値が大きくなっているといえる。よって、指標 B から C への変更においても目的としていた効果をあげることができた。

以上のことから、16 分音符化した音程の列として表したフレーズ間の類似度をあたえる指標としては、三つの指標の中で指標 C が最も有効だといえる。

## 6 楽曲間の類似度

本章では、第 5 章であたえた指標をもとに、楽曲間の類似性指標の定義を試みる。

## 6.1 類似度の定義

二つの楽曲間の類似度を、それぞれの含むフレーズ間の類似度に基づいて定めることを考えよう。一方の楽曲の任意の位置から始まるフレーズは、他方の楽曲において、やはり任意の位置から始まるフレーズと対比させる必要がある。そこで、比較する楽曲間の任意の位置から始まるフレーズのすべての組合せを考えることにする。ここで、フレーズの開始位置は、16分音符単位で変化させることとする。フレーズの長さには制限をおかず、比較する二つのフレーズのうち短い方に合わせることにする。すべてのフレーズの対に対して、それぞれ類似度を求め、その総和を、それぞれの楽曲の長さの積で割ったものを類似度とする。

形式的には、次のようになる。Σ上の長さの等しい文字列間の類似度を与える指標をδとする。任意の文字列  $x, y \in \Sigma^+$  に対して

$$h_\delta(x, y) = \delta(x[1:m], y[1:m])$$

と定める。ここで、 $m = \min(|x|, |y|)$  とする。このとき、文字列  $x, y \in \Sigma^+$  の間の類似度  $\text{SIM}_\delta(x, y)$  を以下で定める。

$$\text{SIM}_\delta(x, y) = \frac{1}{|x| \cdot |y|} \left( \sum_{0 < \ell \leq |x|} h_\delta(x[1+\ell:|x|], y) + \sum_{0 < \ell \leq |y|} h_\delta(x, y[1+\ell:|y|]) \right)$$

## 6.2 評価

変奏曲データ K.25, K.265, K354 を用いて、6.1 節で提案した楽曲間の類似性指標の検証を行なった。Mongeau らの指標との比較結果を表 3 に示す。ただし、提案指標については、フレーズ間の指標に指標 C を  $s = 30, t = 4$  として用いた。

この結果において K.354-4 や K.25-5 に注目していただきたい。Mongeau らの指標では、これらの非類似度の値は比較的小さく、第 2 位と第 4 位に順位づけられている。一方、提案指標では、これらの類似度の値は 0 となる。実際にこれらを聴いて判断したところ、K.265-th に全

表 3: K.265-th に対する結果

順位	Mongeau らの指標		提案指標	
	曲名	非類似度	曲名	類似度
1	K.265-5	28.4	K.265-5	11.2
2	K.354-4	38.5	K.265-9	10.4
3	K.265-12	39.9	K.265-3	3.0
4	K.25-5	42.5	K.265-1	2.9
5	K.25-th	46.8	K.265-11	0.9
6	K.265-9	47.0	K.265-7	0.5
7	K.354-12	47.2	K.354-12	0.3
8	K.265-11	48.9	K.25-3	0.3
9	K.25-7	49.0	K.265-12	0.1
10	K.354-2	54.6	K.25-7	0.0

く似ていなかった。すなわち、この類似度の減少は妥当なものといえる。

次に、K.265-9 に注目しよう。Mongeau らの指標では、比較的非類似度の値が大きく、上で述べた K.25-5 や K.354-4 よりも低い順位となっている。一方、提案指標では、類似度の値は大きく、順位も高い。ここでも、実際にこれらを聴いて判断したところ、K.265-th に非常によく似ている。すなわち、この類似度の増加は妥当であろう。

最後に、K.265-12 をみてみよう。Mongeau らの指標では、第 3 位となっているが、提案指標では、低い類似度の値を示している。これは、同じ K.265 の変奏であるが、3 拍子の変奏であってあまり似ておらず、この結果は妥当なものといえる。

以上のことから、提案指標は、Mongeau らの指標との比較において、有効な指標であるといえる。

## 7 リズムに注目した類似性

本論文では、これまで、主旋律を 16 分音符を単位とした音程の列に変換し、これに対する類似性を考察してきた。しかし、楽曲の類似には、音程の他にも様々な要因が関与している。この章では、リズムの面から見た類似性を扱う。





られたが、変奏曲データではあまりみられなかった。これは、変奏曲は、リズムに変化をつけて「変奏」を作ることが多いという性質を考えると、もっともな結果といえよう。

## 8 おわりに

本論文では、楽曲の主旋律間の類似性指標を提案し、Mongeau らの指標との比較によって、その有効性を示した。また、リズムに注目した場合の指標を提案し、検証を行なった。

今後の課題としては、次のようなものが考えられる。第1に、提案した指標は、編集操作を置換のみに限定した。しかし、適度な挿入、削除の考慮は有効であると考えられ、この操作を導入した場合の指標の検討は必要であろう。

第2に、音程情報の列を入力としたときの指標において、本論文では、一致か不一致かしかみていない。Mongeau らは、不一致の場合を、和声学に基づき細かく分けて扱ったが、それほど細かくするのではなく、協和・不協和・一致といった3段階の基準を採用することで、より有効な類似性指標が得られると考えられる。これについては、ある程度有効性を確かめている [6]。

第3に、本論文では、曲の主旋律という単旋律にのみ注目してきた。一方で、音楽は単旋律で成り立つことは稀であり、ほとんどの場合は和音を基に成り立っているといえる。実際、「似ている」と感じる場合に、和音の影響は大きく、主旋律は、その和音の一部であるともいえよう。したがって、和音を含めた類似度の定義も重要な課題といえる。この問題は、ある時点で同時に鳴っている音の集合を一つの記号とみなすことで、今回の指標を応用することも可能であろう。

第4に、文献 [5] において、和歌の類似度として、データベースにおける文字列の頻度を考慮した指標が提案されている。この方法の楽曲の類似度に対する応用は大変興味深いものである。

最後に、ある程度大規模なデータを用いた類似曲抽出実験である。これは、既に、約1000曲の日本の歌謡曲のデータを所有しているが、計算量の問題から、実験を行なえずにいる。よ

って、より効率的なアルゴリズムの設計が急がれる。

以上のように、様々な課題をあげることができ、いずれもが大変興味深い。本論文で、類似度の定義をする際に得られた主旋律に対する知見は、今後、これらの課題に取り組むにあたって、大いに貢献するであろう。

## 参考文献

- [1] D. Bainbridge. MELDEX: a Web-based melodic locator service. In W.B.Hewlett and E. Selfridge-Field, editors, *Melodic Similarity: concepts, procedures, and applications*, pp. 223-229. The MIT Press, 1997-98.
- [2] A. Ghias, J. Logan, D. Chamberlin, and B. Smith. Query by humming - musical information retrieval in an audio database. In *ACM Multimedia 95 - Electronic Proceedings*, 1995.
- [3] T. Kageyama and Y. Takashima. A melody retrieval method with hummed melody. *Trans. Inst. Electron. Inf. Commun. Eng. D-II*, Vol. J77D-II, No. 8, pp. 1543-1551, 1994.
- [4] M. Mongeau and D. Sankoff. Comparison of musical sequences. *Computers and the Humanities*, Vol. 24, No. 3, pp. 161-175, 1990.
- [5] K. Tamari, M. Yamasaki, T. Kida, M. Takeda, T. Fukuda, and I. Nanri. Discovering poetic allusion in anthologies of classical Japanese poems. In *Proc. 2nd International Conference on Discovery Science (DS'99)*, pp. 128-138, 1999.
- [6] 門田隆史, 石野明, 竹田正幸, 松尾文碩. 音符列比較における類似性指標の評価. 第59回情報処理学会全国大会講演論文集 (2), pp. 17-18, 1999.